RSAC 2025 Conference

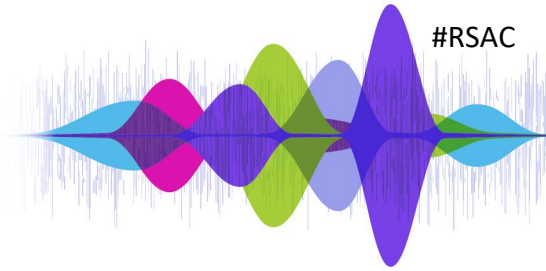
Many Voices.
One Community.

SESSION ID: PART4-W08

How To Use LLMs to Augment Threat Alerts with the MITRE Framework

Nakkul Khuraana
Sr. Data Scientist
OpenText

Dr. Hari Koduvely
Lead Data Scientist
OpenText

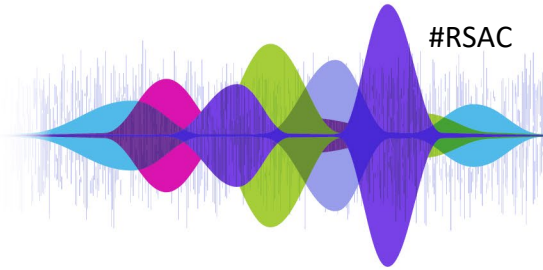


Disclaimer

Presentations are intended for educational purposes only and do not replace independent professional judgment. Statements of fact and opinions expressed are those of the presenters individually and, unless expressly stated to the contrary, are not the opinion or position of RSA Conference LLC or any other co-sponsors. RSA Conference LLC does not endorse or approve, and assumes no responsibility for, the content, accuracy or completeness of the information presented.

Attendees should note that sessions may be audio- or video-recorded and may be published in various media, including print, audio and video formats without further notice. The presentation template and any media capture are subject to copyright protection.

© 2025 RSA Conference LLC or its affiliates. The RSAC and RSAC CONFERENCE logos and other trademarks are proprietary. All rights reserved.



Content

- Introduction
- Overview
- Journey
  - Initial Attempts
  - Enhancements to RAG
- Evaluation and Results
- Case Study
- Conclusion and Takeaways

RSAC 2025 Conference

Introduction

Where are We Failing Today?

DATA VOLUME

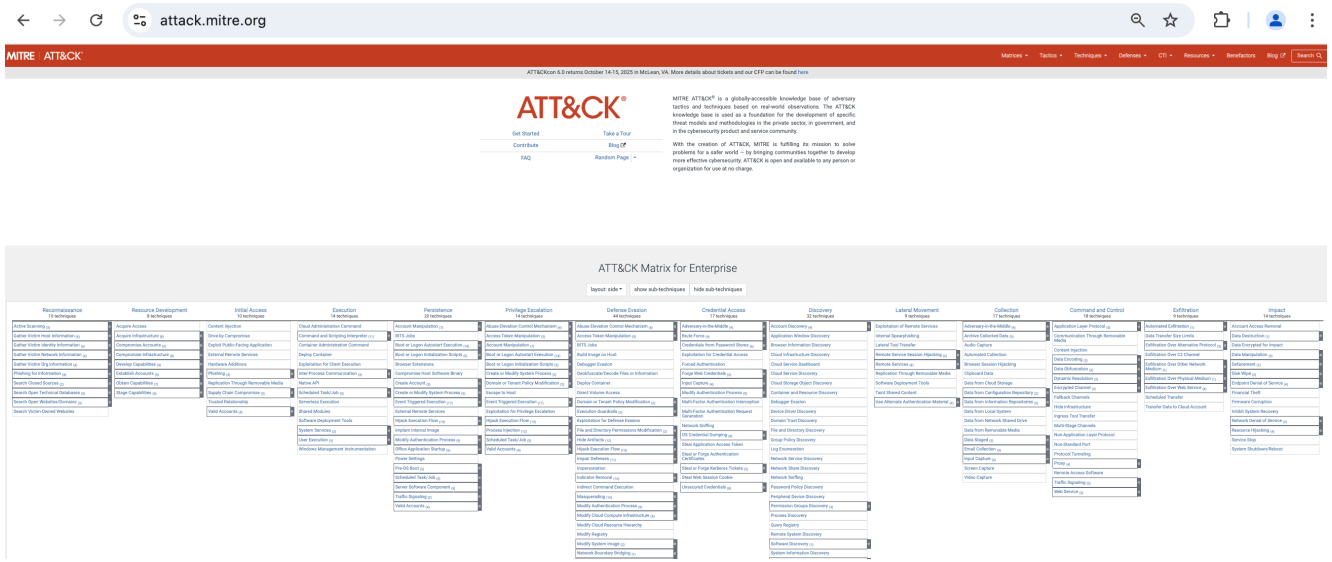
THREAT VELOCITY

CYBERSECURITY TALENT SHORTAGE

The Art of Storytelling

Reference - Shutterstock

CAVE ART



MITRE ATT&CK® FRAMEWORK

MITRE – What's in It?

Exfiltration Over Alternative Protocol

Sub-techniques (3)

ID	Name
T1048.001	Exfiltration Over Symmetric Encrypted Non-C2 Protocol
T1048.002	Exfiltration Over Asymmetric Encrypted Non-C2 Protocol
T1048.003	Exfiltration Over Unencrypted Non-C2 Protocol

Adversaries may steal data by exfiltrating it over a different protocol than that of the existing command and control channel. The data may also be sent to an alternate network location from the main command and control server.

ID: T1048

Sub-techniques: T1048.001, T1048.002, T1048.003

ⓘ Tactic: Exfiltration

ⓘ Platforms: IaaS, Linux, Network, Office Suite, SaaS, Windows, macOS

How Associating Threat Alerts with MITRE Helps the Issue?

By 11:30, another system (10.0.1.12) sent a 1.2GB file, encrypted_archive.zip, to an unfamiliar IP (45.33.32.156).

At 12:10, user j. smith used ftp.exe to transfer customer_data.7z to ftp. evil-exfil[.]com.

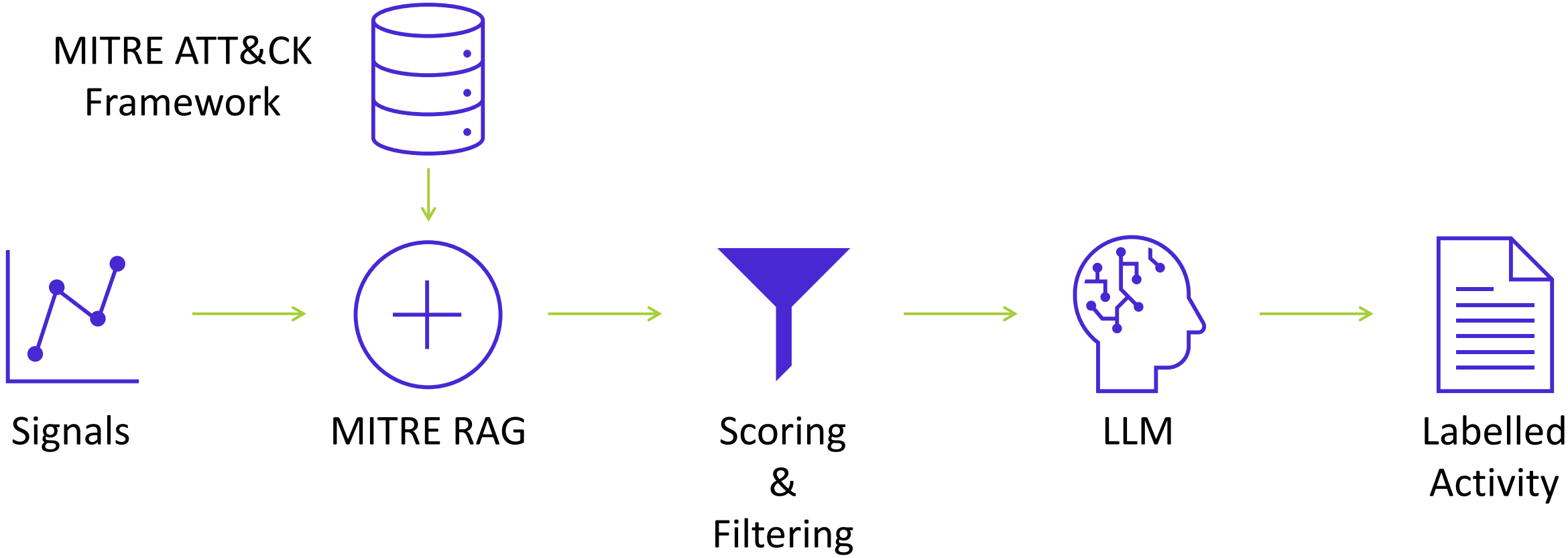
Later, at 13:05, 10.0.3.22 sent a 1024-byte ICMP Echo Request to 89.44.12.77.

"T1048": "Unauthorized FTP upload of compressed sensitive data. ICMP tunneling detected (data embedded in payload)"

RSAC 2025 Conference

Overview

Overview


MITRE ATT&CK Framework
Signals
MITRE RAG
Scoring & Filtering
LLM
Labelled Activity

RSAC 2025 Conference

Journey

Initial Attempts – Manual Annotation

- Manual association of anomalies to MITRE ATT&CK do not scale
  - Number of anomaly models are fixed
  - Ideally each model could be mapped to a MITRE TTP
  - Different attacks could use different processes for the same tactic
  - Large number of anomalies will have to be manually annotated

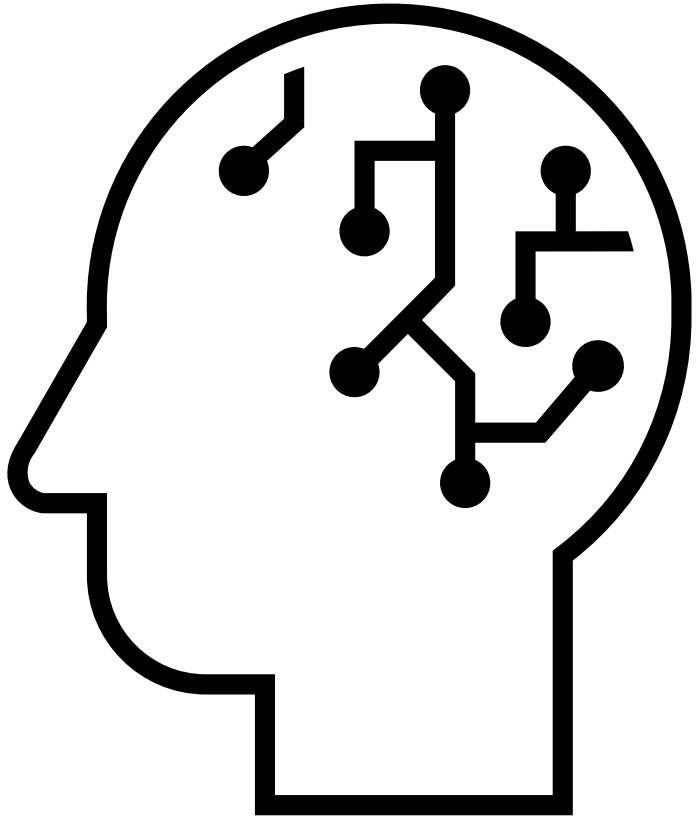
Initial Attempts – ML-based Classification

Threat Type	Technique	Technique Description	Alert ID
Initial Access	T1189	Drive-by Compromise	XX03
Initial Access	T1189	Drive-by Compromise	XX04
Initial Access	T1189	Drive-by Compromise	XX05
Initial Access	T1189	Drive-by Compromise	XX06
Initial Access	T1189	Drive-by Compromise	XX09
Initial Access	T1189	Drive-by Compromise	XX10
Initial Access	T1189	Drive-by Compromise	XX28
Initial Access	T1189	Drive-by Compromise	XX29
Initial Access	T1190	Exploit Public-Facing Application	XX09
Initial Access	T1190	Exploit Public-Facing Application	XX10
Initial Access	T1190	Exploit Public-Facing Application	YY05
Initial Access	T1190	Exploit Public-Facing Application	YY06
Initial Access	T1190	Exploit Public-Facing Application	YY07
Initial Access	T1190	Exploit Public-Facing Application	YY08
Initial Access	T1190	Exploit Public-Facing Application	YY13
Initial Access	T1190	Exploit Public-Facing Application	YY14
Initial Access	T1193	External Remote Services	ZZ02
Initial Access	T1193	External Remote Services	ZZ03
Initial Access	T1193	External Remote Services	ZZ04

- ML classification gives poor results:
  - Lack of good quality labelled data
  - Attack to MITRE TTPs mapping is often one to many

Initial Attempts – Ask an LLM

- Finding associations by directly querying an LLM:
  - Need to use locally hosted small LLMs (e.g. Llama 3.1)
  - They produce hallucinations and non-deterministic results



Initial Attempts – Prompt Tuning

- Improving the prompt
  - Few shot learning examples
  - COT examples
  - Specific instructions for output format

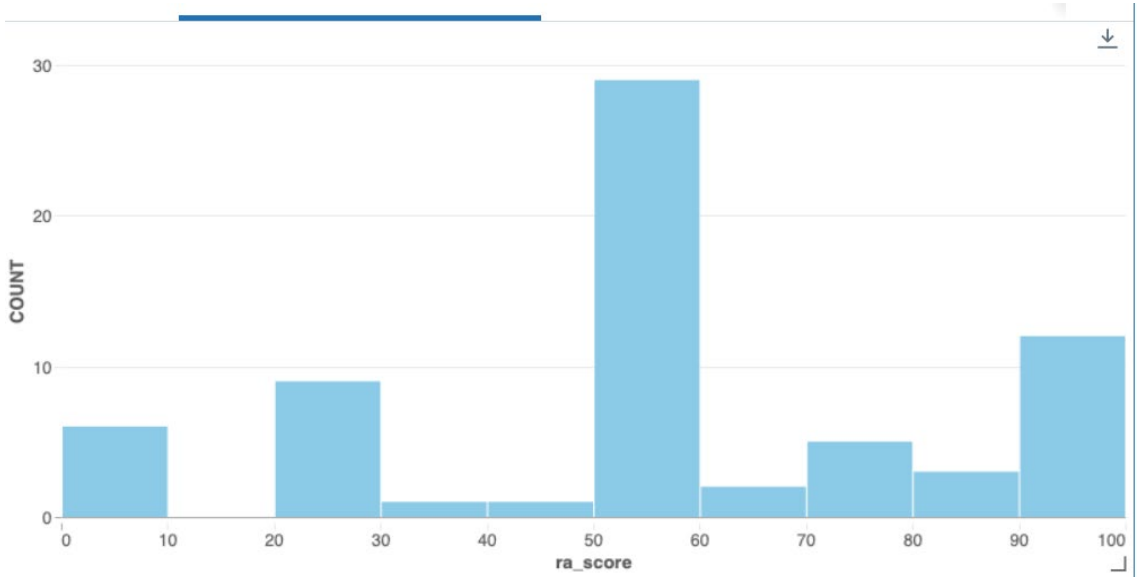
```
prompt = f"""
        <s>[INST]
        You are an expert cyber threat analyst. Your task is to identify the correct MITRE ATT&CK technique involved in the cyber attact described in the field INPUT:

        **Instructions:**
        1. **Output Format**: Only respond with a valid JSON object.
        2. **Valid MITRE Techniques**: Use **only** MITRE techniques from `MATCHING_MITRE`. If a technique is not explicitly in `MATCHING_MITRE`, exclude it.
        3. **Key Formatting:** JSON keys should follow this pattern:
            - `"T#### - Technique Name (Subtechnique Name if applicable)"`
        4. **Correlate Activities Accurately:** If a MITRE technique in `MATCHING_MITRE` is strongly supported by `SUMMARY`, include it with an **explanation**.
        5. **Exclude Weak Correlations:** Do not force weak or unrelated connections.
        6. **Concise and Clear Explanations:** Each technique should have a **brief**, **non-repetitive** explanation (under 300 words).
        7. **Ensure JSON Validity:** Do not include extra text, incorrect characters, or malformed JSON.


        Example1:
        {{
            "date" : "2024-05-12",
            "user" : "user01@domain.com",
            "INPUT": {{"Multiple anomalous activities were detected on May 12th, 2024, related to the user 'user01@domain.com'. These anomalies include multiple failed and successful
            login attempts from various destinations,............."}}
            "MATCHING_MITRE": {{"T1110": {{"technique_name": "Brute Force", "distance": 0.20698959, "subtechniques": ["T1110.001 - Credential Stuffing", "T1110.002 - Password
            Guessing"], "tactics": ["Credential Access"], "score": 0.8030104 }}, "T1021": {{"technique_name": "Remote Services", "distance": 0.41558775, "subtechniques": [],
            "tactics": ["Lateral Movement"], "score": 0.7061641}}, "T1059": {{"technique_name": "Command and Scripting Interpreter", "distance": 0.41474435, "subtechniques": ["T1059.
            003 - Windows Command Shell", "T1059.004 - Unix Shell"], "tactics": ["Execution"], "score": 0.78474435 }}}}
```

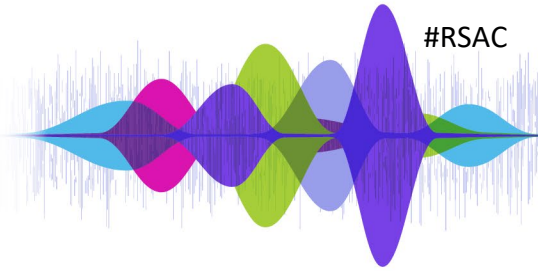
Initial Attempts – Use RAG

- Straightforward RAG also gives sub-optimal results:
  - Retrieval accuracy is only about 50%



Retrieval Accuracy Score Distribution

Enhancements to RAG - Enrichment



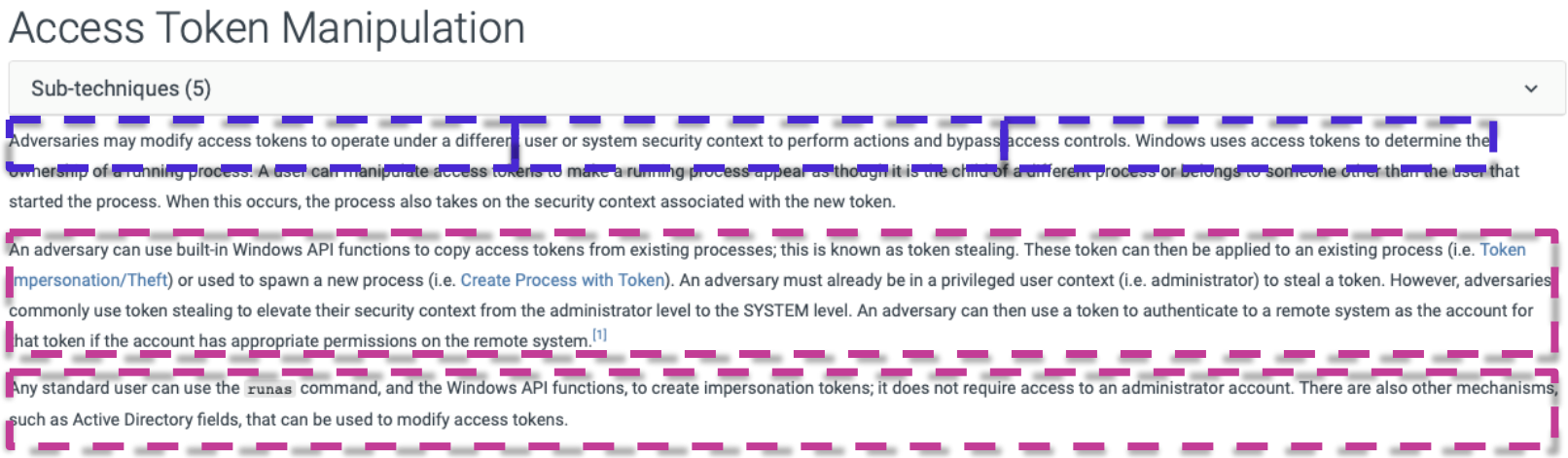
- Pre-processing and enriching the input into the RAG system
  - For each detected risky anomaly, a description is produced using an LLM
  - This is further enriched using more contextual information e.g. details about the processes involved

Enhancements to RAG - Chunking

- Better vector database design to make retrievals more relevant and faster
  - Using *Semantic Chunking* of MITRE documents.
  - Keeping the embedding vectors in memory for faster retrieval.

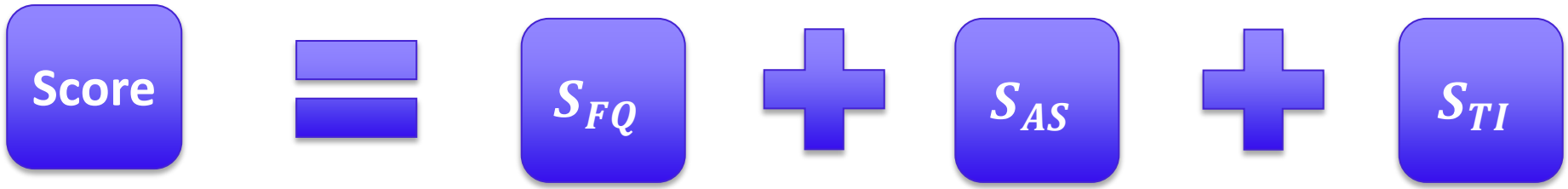
Standard Chunking

Semantic Chunking

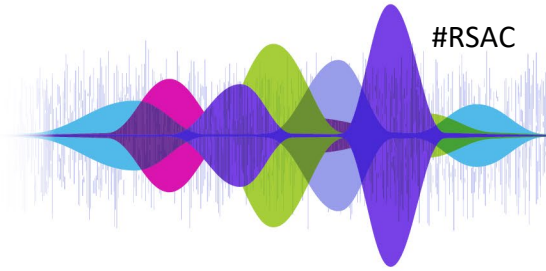

Access Token Manipulation
Sub-techniques (5)
Adversaries may modify access tokens to operate under a different user or system security context to perform actions and bypass access controls. Windows uses access tokens to determine the ownership of a running process. A user can manipulate access tokens to make a running process appear as though it is the child of a different process or belongs to someone other than the user that started the process. When this occurs, the process also takes on the security context associated with the new token.
An adversary can use built-in Windows API functions to copy access tokens from existing processes; this is known as token stealing. These token can then be applied to an existing process (i.e. Token Impersonation/Theft) or used to spawn a new process (i.e. Create Process with Token). An adversary must already be in a privileged user context (i.e. administrator) to steal a token. However, adversaries commonly use token stealing to elevate their security context from the administrator level to the SYSTEM level. An adversary can then use a token to authenticate to a remote system as the account for that token if the account has appropriate permissions on the remote system.[1]
Any standard user can use the runas command, and the Windows API functions, to create impersonation tokens; it does not require access to an administrator account. There are also other mechanisms, such as Active Directory fields, that can be used to modify access tokens.

Enhancements to RAG - Scoring

- Retrieval at each anomaly level and aggregation of the results to surface the top N likelihood MITRE techniques.
  - Each of the retrieved result is scored using a scoring function.
  - Score consists of 3 components.



Enhancements to RAG – Frequency Dependent Score Component



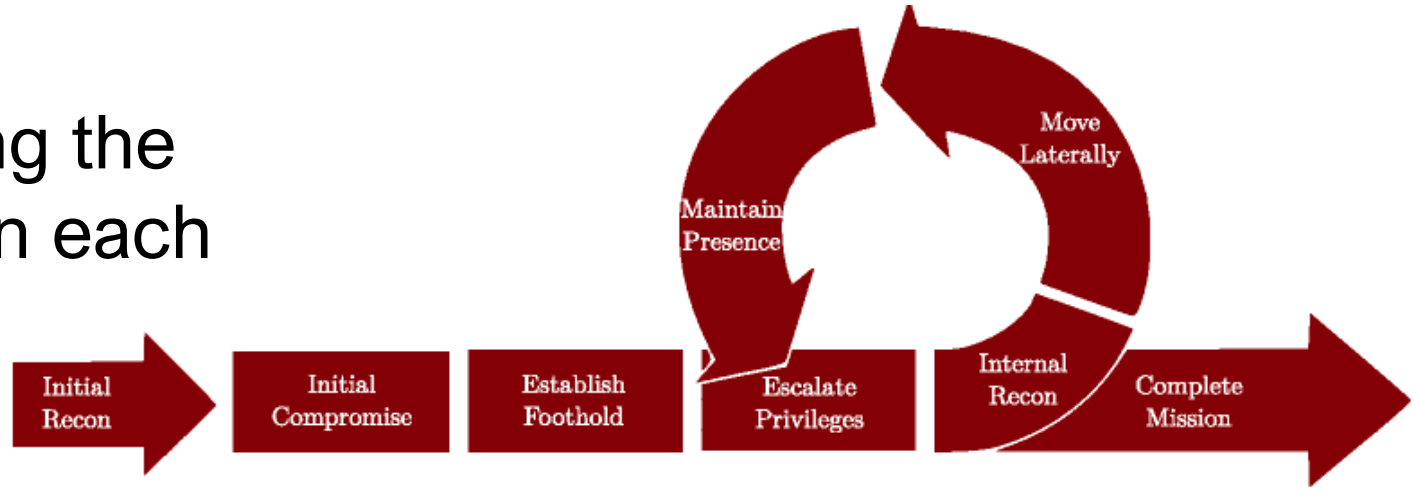
For each MITRE technique:

- Compute daily aggregates:
  - Count of technique in the process anomalies list C_P
  - Count of technique in the RAG retrieved results C_R
- Compute the distance statistics:
  - Average distance D_{Avg}
  - Maximum distance D_{Max}
- Compute
  - Total anomalies N_A
  - Total process anomalies N_P

$$S_{FQ} = \frac{N_P}{N_A} C_P \, f(D_{Avg}, D_{Max}, C_R)$$

Enhancements to RAG – Attack Stage Dependent Score Component

- A weight depends on which stage the attack is
- Higher the weight for more advanced stage
- Weight is computed using the volume of attacks seen in each stage

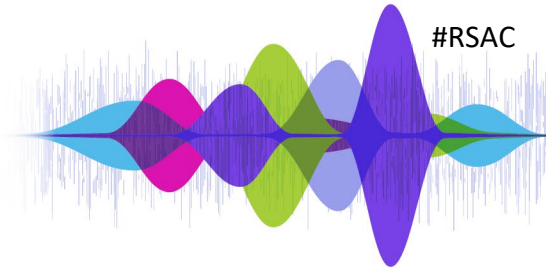


Reference - Mandiant Attack Cycle Publication

Enhancements to RAG – Threat Intelligence-based Score Component

- A weight depends on whether the elements of the attack is found to be associated with TTPs in threat intelligent reports.
- Higher the weight if the attack is seen in more reports.
- Higher the weight if the report is more recent.

Enhancements to RAG – Reasoning and Explanation

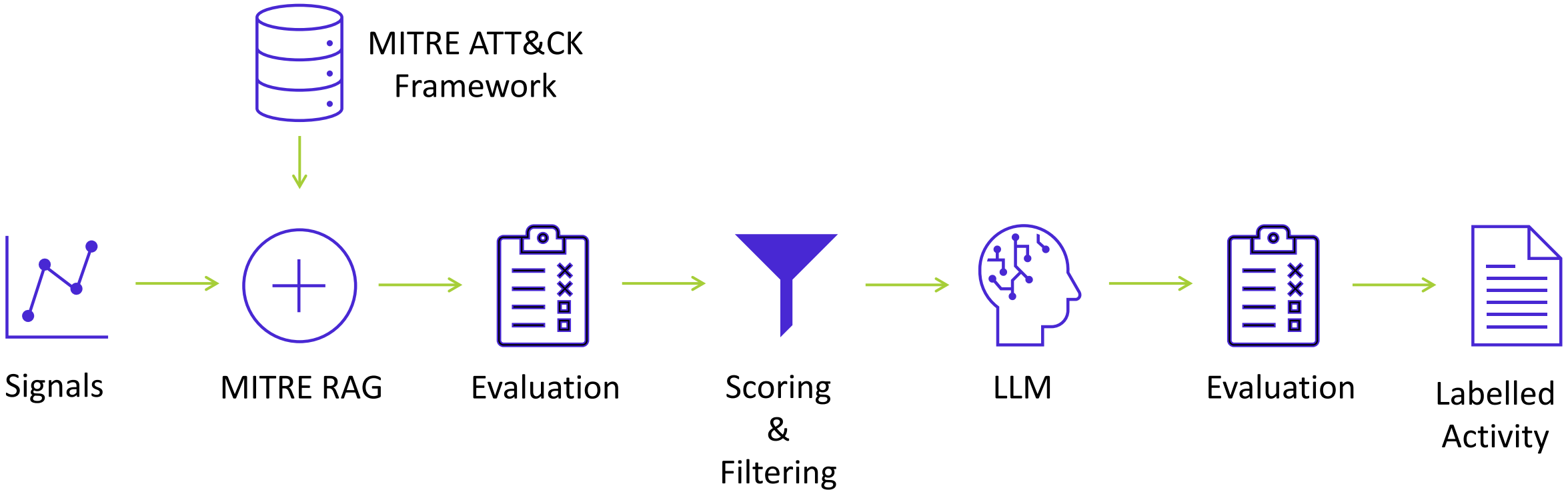


- Reasoning using an LLM to arrive at an explanation of the association
  - T1110 - Brute Force: The unusual number of days with login successes for the user on specific servers and across multiple servers could indicate a brute force attack or a compromised account.
  - T1021 - Command and Control: The execution of the process eot2s03lk2ls.exe, which is rarely seen for this user and has a high importance, could be indicative of command and control activities.

RSAC™ 2025 Conference

Evaluation and Results

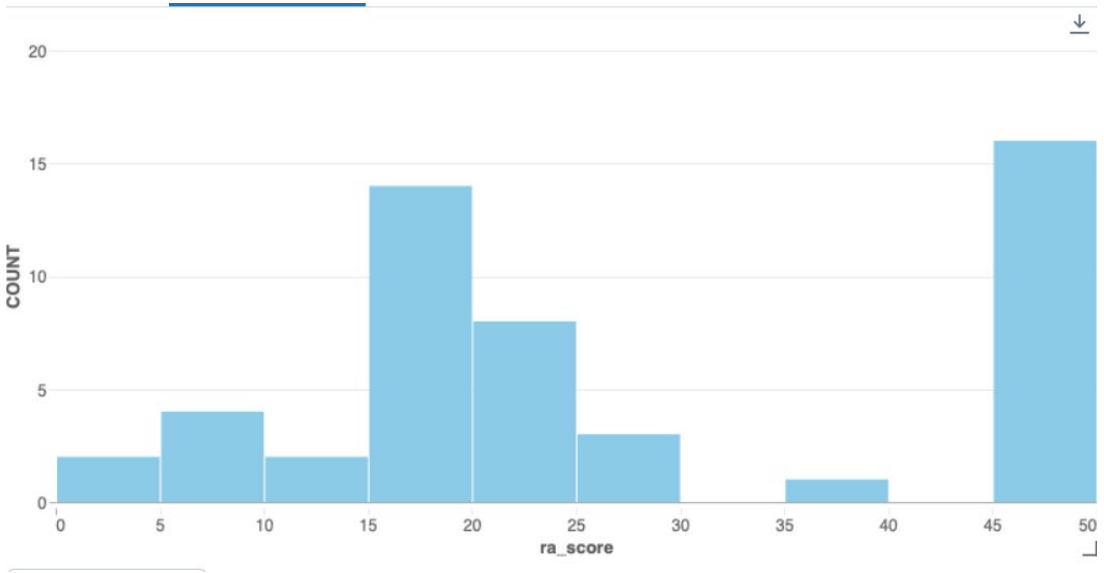
Overview + Evaluation



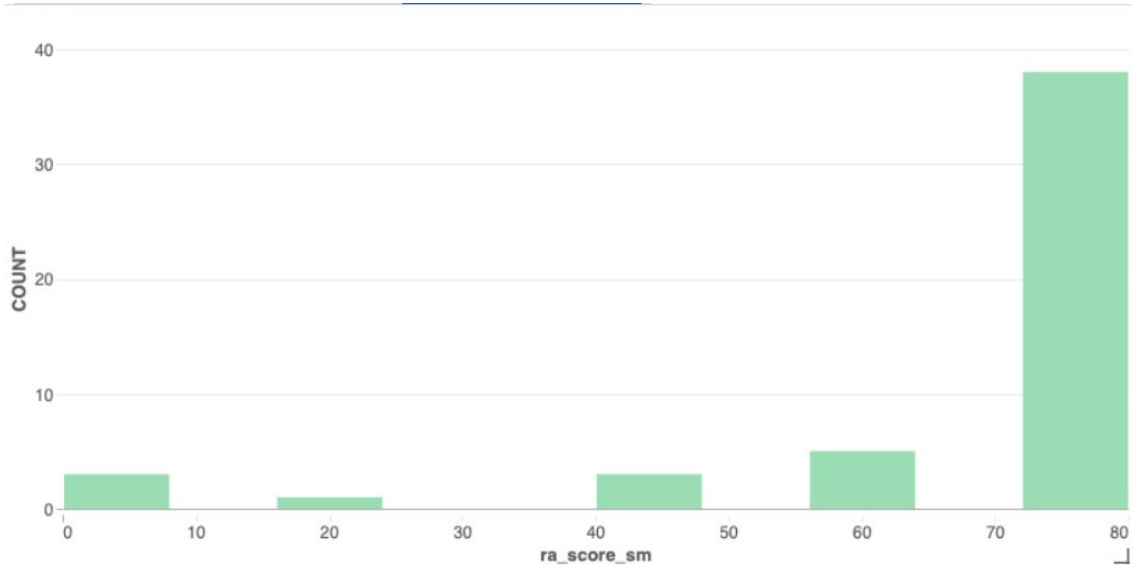
Evaluation of the MITRE ATTACK Matching

- No ground truth data
- Use a Judge LLM for evaluation
- Metrics:
  - Accuracy of MITRE Matching

Accuracy of MITRE Matching



Accuracy with only using RAG



Accuracy with RAG + Scoring

RSAC 2025 Conference

Case study

Selected Signals


```
{RISK_CONTRIBUTION=1.717190384327542, ANOMALY_DESCRIPTION=OWN,
Process=net.exe, User=svc_service_booking@contsocorp.com, TIMESTAMP=2023-
10-10 13:00:00.000},
{RISK_CONTRIBUTION=1.717190384327542, ANOMALY_DESCRIPTION=OWN,
Process=net.exe, User=svc_service_booking@contsocorp.com, TIMESTAMP=2023-
10-10 13:00:00.000},
{RISK_CONTRIBUTION=1.6702337229649875, ANOMALY_DESCRIPTION=OWN,
Process=reg.exe, User=svc_service_booking@contsocorp.com, TIMESTAMP=2023-
10-10 13:00:00.000},
{RISK_CONTRIBUTION=1.6702337229649875, ANOMALY_DESCRIPTION=OWN,
Process=reg.exe, User=svc_service_booking@contsocorp.com, TIMESTAMP=2023-
10-10 13:00:00.000},
{RISK_CONTRIBUTION=1.0000693171203765,
ANOMALY_DESCRIPTION=CorrolationRuleSrv,
User=svc_service_booking@contsocorp.com, TIMESTAMP=2023-10-10
13:00:00.000},
{RISK_CONTRIBUTION=1.0000693171203765, ANOMALY_DESCRIPTION=OWN,
Process=eot2s03lk2ls.exe, User=svc_service_booking@contsocorp.com,
TIMESTAMP=2023-10-10 13:00:00.000},
{RISK_CONTRIBUTION=1.0000693171203765, ANOMALY_DESCRIPTION=OWN,
Process=eot2s03lk2ls.exe, User=svc_service_booking@contsocorp.com,
TIMESTAMP=2023-10-10 13:00:00.000},
{RISK_CONTRIBUTION=1.000069316654241,
ANOMALY_DESCRIPTION=4624_~success~|~10~|~default~,User=svc_service_booking@
contsocorp.com, Server=ALL_SERVERS, TIMESTAMP=2023-10-10 13:00:00.000}
```

MITRE ATT&CK Vector Database

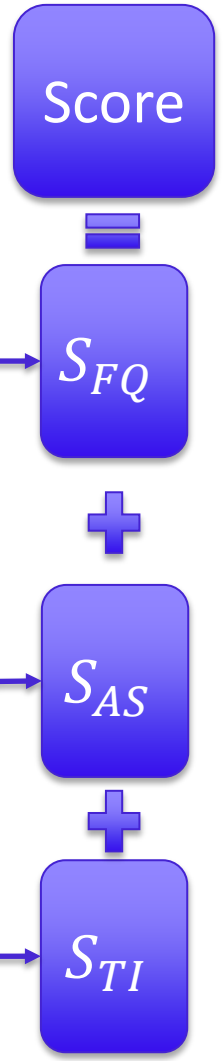

```
{"ID": "T1574.004", "name": "Hijack Execution Flow: Dylib Hijacking",
"url": "https://attack.mitre.org/techniques/T1574/004", "tactics": "Defense
Evasion, Persistence, Privilege Escalation", "chunk_no":} "4817160198"},
{"ID": "T1602", "name": "Data from Configuration Repository", "url":
"https://attack.mitre.org/techniques/T1602", "tactics": "Collection",
"chunk_no": "4869880978"} Add of existing embedding ID: {"ID": "T1574.004",
"name": "Hijack Execution Flow: Dylib Hijacking", "url":
"https://attack.mitre.org/techniques/T1574/004", "tactics": "Defense
Evasion, Persistence, Privilege Escalation", "chunk_no": "8921924710"} Add
of existing embedding ID: {"ID": "T1602", "name": "Data from Configuration
Repository", "url": "https://attack.mitre.org/techniques/T1602", "tactics":
"Collection", "chunk_no": "1206942336"}
```

Selecting Top-N MITRE Techniques

Technique	Tactic	Name	$Score_{FQ}$
T1110	Credential Access	Brute Force : Password Guessing	.80
T1110	Credential Access	Brute Force : Password Guessing	.68
T1021	Lateral Movement	Remote Services : Cloud Services	.75
T1021	Lateral Movement	Remote Services : Cloud Servies	.63
T1606	Credential Access	Forge Web Credential	.51

Technique	$Score_{AS}$
T1110	.75
T1021	.60

Technique	Severity	Description	$Score_{TI}$
T1021	INFO	Logged in via cloud service	.75
T1110	INFO	Brute Force : password guessing	.50
T1129	INFO	Link function at runtime on Linux	.10



Labelled Activity with MITRE Techniques

svc_service_booking@contsocorp.com - Unusual Login Activity and Unusual Processes"

"summary”: "Based on the data, the user svc_service_booking@contsocorp.com exhibited several anomalous activities on October 10th, 2023. These anomalies include an unusual number of days with login successes for the user on specific servers, such as ad.0905.can.com, as well as an unusual number of days with login successes across multiple servers. Additionally, the user executed several unusual processes, including hikaru.exe, reg.exe with high importance, and eot2s03lk2ls.exe, which is rarely seen for this user and has a high importance.

The MITRE attack framework for this anomaly would be:

T1110 - Brute Force: The unusual number of days with login successes for the user on specific servers and across multiple servers could indicate a brute force attack or a compromised account.

T1021 - Command and Control: The execution of the process eot2s03lk2ls.exe, which is rarely seen for this user and has a high importance, could be indicative of command and control activities.

RSAC 2025 Conference

Conclusion and Takeaways

Many Voices.
One Community.

Conclusion

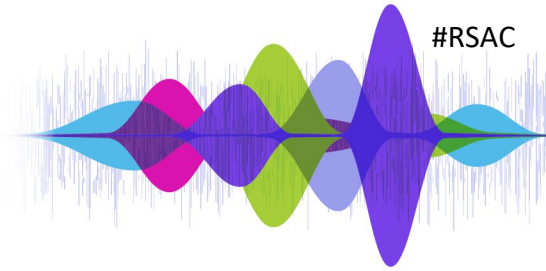
Automatic Association of MITRE

Learn from the Journey

- LLM is not a silver bullet
- Experiment and Iterate

Do it Right

Takeaways



In 3-5 Days

- Select the Right Signals
- Design Meaningful Aggregation
- Use Gen AI Tools

In 1-3 Weeks

- Optimize Your Prompt

In 1-2 Months

- Automate It
- Productize It

RSAC™ 2025 Conference

Many Voices.
One Community.

Thank You

Visit Us at Booth

N-4535